

浅谈图书馆数据备份自动化

洪丹萍

(厦门大学图书馆, 福建厦门 361005)

摘要 图书馆自动化系统的核心是数据。为了保证系统的安全运行, 数据备份是必要的、也是重要的。本文阐述了如何在网络环境下实现数据备份自动化, 以及备份自动化的必要性和可行性。

关键词 网络; 图书馆; 数据; 自动备份

中图分类号: TP309.3 文献标识码: B 文章编号: 1008-7826 (2002) 04-0101-03

随着计算机技术、计算机网络的高速发展与图书馆自动化系统软件的不断完善, 大部分图书馆都进入了办公自动化阶段。当我们深切感受到计算机网络管理的优越性的时候, 也越来越依赖自动化系统。然而这种依赖是令人担忧的。因为事实上, 不论是从硬件、软件、网络环境等方面来看, 计算机网络系统都不是十分可靠的, 尤其是在 Internet 日益普及, 系统开放性日益加大的今天, 这种担忧是非常必要的。

图书馆计算机网络系统是一个庞大复杂的系统, 如果缺乏严格的科学管理, 缺乏紧急情况的应急措施, 那么对系统的安全性的潜在危害会远远大于软、硬件。这是整个系统不安全因素中最重要的因素, 因为再完善、再安全的硬件系统, 都有可能遭到人为的有意或无意的破坏。

1 数据备份的必要性与重要性

数据是图书馆最宝贵的财富和资源, 是图书馆的核心。图书馆中的一切活动几乎都以这些数据为基础。然而, 一旦数据遭到破坏和丢失, 将影响整个图书馆业务工作的正常开展, 并可能造成无法挽回的重大损失。在图书馆自动化系统中, 数据有两大特点, 一是数据量大且集中, 二是具有变化性和流动性。因而决定了系统具有较大的风险性, 任何一个环节失误, 都可能造成不堪设想的后果。系统中馆藏信息、采访部门的订购信息、编目部门的在编书目数据、瞬息万变的读者借还图书的信息以及工作人员的工作日志等等, 假如没有做好数据的备份工作, 一旦发生不测, 那么就没有任何根据来证明读者是否有借或还书, 到底借了多少本书; 采访、编目部门工作人员辛苦输入的数据也会丢失, 图书馆的工作也就无法正常进行, 带来的损失将是无法估计。

不论数据库管理系统提供了多么严格的防范措施, 数据库遭破坏的事件还是有可能发生的, 如人为的入侵, 即非法用户对网络系统、自动化系统的非法使用与合法用户的越权使用所造成的数据被非法窃取、篡改和破坏; 或系统突然发生故障, 导致数据库的破坏。可以想象, 如果一个馆的数据库丢失或被破坏, 那将是多么严重的问题, 会给图书馆造成巨大的, 不可挽回的损失。为了避免难以弥补的损失, 唯一的途径就是对系统进行有效、可靠的备份。也就是说, 做好有效、可靠的数据备份是必须的, 也是必不可少的。但是, 如何做好数据备份工作呢?

2 数据备份的策略

图书馆数据的备份可以采用完全备份和增量备份策略。

我校图书馆采用深圳图书馆研制开发的图书馆自动化管理系统——ILASII,该系统包含菜单文件、屏幕文件、日志文件和数据库等。一般情况下,采访、编目、流通等部门的日常工作都会随时改变其相关的 database 数据库,而菜单文件、屏幕文件等一般没变化。另外,在 database 数据库中,外部库(中、西文外部书目库和中西文期刊外部书目库)是编目部工作人员用来接收外来数据,一般隔一段时间才接收一批数据。由于外部库特别大而且平时没什么变化,没必要每天备份。因此,日常备份不一定都要完整地把整个图书馆自动化管理系统全部备份,而只需备份变化大的,记载采访、编目工作人员以及流通部门读者借还图书记录的数据库。也就是只备份除外部库外的 database 数据库。这样的备份就叫做增量备份。同时每隔一段时间进行一次完全备份,即备份整个图书馆自动化管理系统,包括菜单文件、屏幕文件、日志文件等。考虑到图书馆的数据库都非常庞大,若在图书馆工作高峰期间备份,往往要与系统的其他运作争夺资源,达不到备份数据的效果,反而影响系统的其他部门正常工作。根据我馆的实际运作情况,我们采用每周一次完全备份,每天两次增量备份的策略。把备份时间安排在中午和晚上,采访、编目工作人员下班且流通量少的时间进行数据库备份。

3 数据备份的方法

一位专家曾经这样描述备份工作,“备份的意思就是每天你都不要忘了从磁带机中取出旧磁带,放入新磁带贴上标签,标上号码,最后还不要忘了定期清洗磁头。。。”这是描述早几年图书馆备份数据的情况,用这种方法备份很容易出错,例如弄混了磁带的序号,不小心插错了磁带、忘记了清洗磁头,甚至忘记了备份等等。这些情况一旦发生,就会给系统中的数据带来严重的潜在的威胁。

随着计算机硬盘容量的不断增大,图书馆数据的备份工作改为硬盘备份,然后用人工把备份数据传输到另一台服务器。这种完全程序化的、十分单调的操作过程毫无趣味性可言,令人望而生厌。因此很多系统管理人员都患有“备份厌烦症”。而且在传输过程中,有时会出现网络堵塞,传输速度慢,工作人员疲劳或其他种种原因而造成数据传输不成功。由于人工介入后总免不了要出错,因此,最理想的备份方法就是完全不需要人工的介入,或将人工的介入降到最低的程度,为了达到这个目的,必须采用备份自动化方法。

3.1 数据备份自动化

目前,大多数图书馆都实现图书管理自动化,而自动化系统的操作平台大都是 UNIX 系统。在 UNIX 系统中,提供 cron 命令。cron 是一个作业调度系统,它可以用来以规定好的间隔定期执行预定的命令。通常情况下,cron 执行的命令都是批处理程序,用来完成日常琐事。因此,要让系统做好数据备份自动化,要做两项准备工作:

(1) 先写好批处理程序,即分别写好中午、晚上等数据备份批处理程序,如中午数据备份的批处理程序名为 backup.am,晚上数据备份的批处理程序名为 backup.pm,而完全备份的批处理程序名为 backup.Fri 等。

(2) 设置好 cron 作业,让 cron 中午执行 backup.am 批处理程序,晚上执行 backup.pm 批处理程序等,把数据备份到另一目录。

要备份完整的数据,最好在流通量少,采访、编目数据基本不变的时间段进行数据备份,每天两次的增量备份安排在中午 12:30 和晚上 22:30。每周一次的完全备份也是在晚上、工

作人员都退出图书自动化管理系统后进行。

工作人员有可能在工作中由于误操作或某种原因出错，要求恢复数据，因此，最好把备份数据打包，并打上备份日期，这样有利于备份数据的管理，随时为恢复数据做好准备工作。另外，为了给系统节省空间，可以把打包的备份数据压缩。因为压缩后的备份数据文件可以节省2/3空间。

3.2 数据传输自动化

在本机做数据备份，虽然备份目录与正在运行的数据库目录不同，但万一主服务器出现故障，或由于黑客的攻击等，主服务器瘫痪了，即使你做了完善的备份工作，你所做工作的也等于白做。因此，必须有一台备用服务器。最好的方法是在备用服务器上安装与主服务器相同的系统以及图书馆自动化软件、设置相同的运行环境。把备份数据传输到这台备用服务器上。这样，万一主服务器发生意外，维修可能不是一天、两天就可以完成，如果有一台备用服务器，把备份数据直接传输到备用服务器上，一旦主服务器出现故障，就可以马上启用备用服务器，只要把标明相应日期的打包压缩数据备份文件解压、拆包，备用服务器就可以马上运行。使图书馆的工作保持正常，把损失降低到最低限度。

要如何把备份数据从主服务器传输到备用服务器呢？可以用FTP软件把备份数据传输到备用服务器。这种方法需要人工操作。人工操作存在着管理员在传输过程中出错的可能性，或传输备份数据不及时，或根本把这项任务给忘了。要及时把备份数据传输到备用服务器，管理员必须每天两次（中午和晚上）做传输备份数据工作，而每天中午、晚上让管理员做传输备份数据工作也太单调、太辛苦了，也很容易出错。因此，最好的方法还是应该让系统自动传输，而不需要人工的干预。

要做到让系统定时自动传输备份数据，必须做以下几个步骤的工作。

（1）在备用服务器先建一用户，专门用来传输备份数据的用户。为了安全起见，必须设置登录密码，并指定登录路径，即登录路径为备份目录。

（2）主服务器根目录编写.netrc文件。

在UNIX系统中，.netrc文件包含通过网络登录到远程主机进行ftp文件传输的数据。这个文件驻留在系统的根目录里。假如备用服务器IP为210.12.52.48，该传输备份数据的用户名为usr2；密码为usr2abc；指定登录路径为/usr2/database.backup；主服务器的备份路径为/usr1/database.backup。则.netrc格式如下：

（3）设置主服务器的cron作业，待数据备份后直接自动传输到备用服务器。

这样，数据从备份、打包、压缩、打上备份日期到把备份数据传输到另一台服务器都是让系统自己完成，不用人工参与，真正实现备份、传输数据自动化。

machine	210.12.52.48	login	usr2	password
usr2abc				
macdef	init			
binary				
lcd	/usr1/database.backup			
mput	*			
bye				
quit				

参考文献

- [1] 廖利娟. 高校图书馆数据备份调查分析[J]. 图书馆建设, 2000 (1)
- [2] 杨晓红. 数据备份的误区及备份自动化解决方案[J]. 重庆师范学院学报 (自然科学版), 2000: 17 (1)
- [3] 向桂林. 谈谈ILAS系统的数据备份[J]. ILAS通讯, 2000: (1)